

基于大样本土壤光谱数据库的氮含量反演

王乾龙¹ 李 硕¹ 卢艳丽² 彭 杰³ 史 舟^{1,4} 周炼清^{1*}

¹ 浙江大学环境与资源学院农业遥感与信息技术应用研究所, 浙江 杭州 310058

² 中国农业科学院农业资源与农业区划研究所农业部作物营养与施肥重点开放实验室, 北京 100081

³ 塔里木大学植物科学学院, 新疆 阿拉尔 843300

⁴ 浙江大学唐仲英传感材料及应用研究中心, 浙江 杭州 310058

摘要 充分挖掘大样本土壤光谱库中有效信息, 建立普适性强的土壤全氮(TN)含量反演模型, 是高光谱分析的重要应用方向之一。研究采用偏最小二乘回归(PLSR)全局建模、局部加权回归(LWR)和模糊 K 均值聚类结合 PLSR(FKMC-PLSR)局部建模三种方法, 分别建立了来自中国西藏、新疆、黑龙江、海南等 13 个省采集的 17 种土类 1661 个土壤样本 TN 值的高光谱反演模型, 并对浙江省 104 个水稻土样本进行模型验证。结果表明, 在大样本下 PLSR 全局模型对高 TN 值待预测样本存在低估现象, 导致整体预测精度偏低; LWR 和 FKMC-PLSR 局部模型比 PLSR 全局模型能够更为准确地反演 TN 含量。研究结果可为利用大样本光谱数据库建立稳定性和普适性较高的土壤 TN 含量预测模型提供参考。

关键词 光谱学; 土壤光谱库; 局部加权回归; 模糊 K 均值聚类; 土壤全氮; 大样本

中图分类号 TP7; S158.2 **文献标识码** A **doi:** 10.3788/AOS201434.0930003

Nitrogen Content Inversion Based on Large Sample Soil Spectral Library

Wang Qianlong¹ Li Shuo¹ Lu Yanli² Peng Jie³ Shi Zhou^{1,4} Zhou Lianqing^{1*}

¹ Institute of Agricultural Remote Sensing and Information Technology Application,

College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China

² Ministry of Agriculture Key Laboratory of Crop Nutrition and Fertilization, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

³ College of Plant Science, Tarim University, Alar, Xinjiang 843300, China

⁴ Cyrus Tang Center for Sensor Materials and Application, Zhejiang University, Hangzhou, Zhejiang 310058, China

Abstract Building universal deduction models for predicting the soil total nitrogen (TN) content by using data mining of large soil spectral libraries is one of the most important applications of hyperspectral analysis. In this study, 1661 soil samples representing 17 soil types from 13 provinces of China (e. g., Tibet, Xinjiang, Heilongjiang and Hainan) are employed for modeling the soil TN content using global partial least squares regression (PLSR), locally weighted regression (LWR) and fuzzy K-means clustering combined with PLSR (FKMC-PLSR). Another 104 paddy soil samples collected from Zhejiang Province are used to validate the established models. Results showed that when predicting soil TN from a large dataset, global PLSR underestimates high values of TN, which generates an overall low prediction accuracy. By contrast, LWR and FKMC-PLSR perform better than global PLSR. It is suggested that the results can provide useful information for establishing robust and universal models for soil TN prediction using

收稿日期: 2014-03-20; **收到修改稿日期:** 2014-04-30

基金项目: 国家 863 计划(2013AA102301)、国家自然科学基金(41271234)、教育部新世纪优秀人才支持计划(NCET-10-0694)

作者简介: 王乾龙(1988—), 男, 硕士研究生, 主要从事土壤属性高光谱定量分析方面的研究。

E-mail: wangqianlong@zju.edu.cn

导师简介: 史 舟(1970—), 男, 教授, 主要从事土壤高光谱遥感、地统计等方面的研究。E-mail: shizhou@zju.edu.cn

* **通信联系人.** E-mail: lianqing@zju.edu.cn

本文电子版彩色效果请详见中国光学期刊网 www.opticsjournal.net

large soil spectral libraries.

Key words spectroscopy; soil spectral library; locally weighted regression; fuzzy K-means clustering; soil total nitrogen; large sample

OCIS codes 300.6340; 300.6490; 300.6550; 070.4790; 120.5700

1 引 言

可见-近红外(Vis-NIR)漫反射光谱分析技术可广泛应用于地物属性的无损测定,并在土壤属性实验室分析估测中获得成功应用^[1]。与传统化学方法相比,光谱定量分析技术快速、无损、低成本且不会对环境造成污染^[2-4]。基于相似土壤组分和光谱特征的研究建立了适合局部区域的预测模型,充分肯定了其在土壤有机质、氮素、有机碳、水分、粘土矿物中的预测能力^[5-7],但由于各区域土类及成土母质存在较大差异,局部区域预测模型难以外推至其他区域其他土类。同时样本复杂的数据集将导致土壤属性和光谱数据的非线性和空间依赖性增强^[8]。此外,局部模型重复采样浪费资源且共享性差。构建全国乃至全球大样本的土壤光谱库为解决以上问题开辟了新的途径^[9-11]。

国际上最近多次开展过利用国家尺度或全球尺度土壤光谱库建模预测土壤氮素等关键属性的研究^[12-14]。2013年,Stevens等^[15]利用来自于23个国家约20000个样本的欧洲大样本光谱库信息,采用偏最小乘回归(PLSR)、增强回归树(BRT)、Cubist、随机森林和支持向量机方法建立模型对研究区域的农田、草地、林地等土壤的氮、有机碳、阳离子交换量(CEC)、pH和粘土矿物等多种属性进行预测;Peng等^[16]也在2013年利用来自丹麦全国2688个土壤样品建立的光谱库,通过全局、全局结合当地30个已知样本(Spiking方法)和不同方式挑选的三种建模子集,分别进行PLSR建模预测局部地区35个样本的有机碳含量。

近年来,中国陆续开展了利用土壤光谱预测氮含量的研究,但主要是针对局部地区小样本的探索及应用。徐永明等^[17]采用逐步回归方法,选取天津70个样本,研究探讨全氮(TN)含量与光谱反射率的相关关系;卢艳丽等^[18]基于主成分回归和光谱波段特征值对110个黑土样本建立预测模型;李硕等^[19]选取湖北省48个样本,比较分析了以PLSR为代表的线性模型和以BPNN为代表的非线性模型预测土壤TN含量的情况;张娟娟等^[20-21]研究土壤TN估测模型,选取331个样本建模,129个样本验证。目前国内鲜见收集全国不同区域不同土壤类

型以及构建土壤光谱库来验证土壤光谱建模及预测能力的相关研究。建立基于大样本土壤光谱库的普适性模型在野外速测及大范围监测等方面具有重要意义,在高光谱遥感方面,统一模型的建立减少了由于地区不同土类不同而频繁更换模型带来的新的误差和时间的延误^[22]。

本文基于大样本的土壤 Vis-NIR 光谱库,研究如何利用该库建立具有稳定性和普适性的高精度回归模型,来预测某区域未知样本的 TN 含量。

2 材料与方 法

2.1 样品的采集与制备

全国土壤 Vis-NIR 光谱库中的 1661 个样本来自西藏、新疆、四川等 13 个省共计 17 种土类,主要有黑土、水稻土、紫色土、潮土和褐土等;104 个待测土样为浙江省富阳上虞龙游地区的水稻土。所有土样采集深度为表层 0~20 cm,风干、研磨后过直径 2 mm 筛,再经四分法分成两份,分别用于化学分析和光谱测定。表 1 为供试土壤样本 TN 含量的基本统计特征,17 种土类根据 TN 均值含量从高到低排序,min 为最小 TN 值,max 为最大 TN 值,mean 为 TN 均值,SD 为 TN 标准偏差。从表 1 可知,库中样本 TN 含量变幅较大,从 0.016% 到 0.461%,信息量丰富,为全面验证土壤 TN 反射光谱反演模型优劣提供了较完整的样本。

2.2 光谱测定及预处理

供试土样的光谱均使用光纤光谱仪采集。仪器为美国 ASD 公司 FieldSpec 4 型,配备一个 50 W 卤素灯作光源,一根 25°视场角的光纤探头。该仪器由 3 个传感器组成,分别为 UV-VNIR(350~1100 nm),SWIR1(1000~1800 nm)和 SWIR2(1700~2500 nm)。其波长范围涵盖 350~2500 nm,采样间隔为 1.4 nm(350~1000 nm)和 2 nm(1000~2500 nm),重采样间隔至 1 nm,共 2051 个波段。

将土样填满于直径 10 cm、深 1.5 cm 的盛样器皿内,用直尺将表面推平。在测定前,为减弱仪器暗电流和灯源对光谱质量的影响,分别进行暗校正和白板校正。经多次实验最终确定光源距样品 70 cm,光源入射角为 60°,光纤探头垂直距土样表面 10 cm。光谱

表 1 不同土类土壤全氮含量的统计特征
 Tabel 1 Descriptive statistics of soil TN content of different soil types

Soil group	Province	Sample	TN / %			
			Min	Max	Mean	SD
Albic soils	Heilongjiang	32	0.104	0.461	0.211	0.078
Meadow soils	Heilongjiang/Jilin	118	0.067	0.442	0.171	0.067
Black soils	Heilongjiang/Jilin/Henan	167	0.056	0.423	0.171	0.067
Brown soils	Heilongjiang/Liaoning/Jilin	55	0.03	0.376	0.131	0.064
Chernozems	Heilongjiang/Jilin	236	0.043	0.408	0.128	0.048
Aeolian soils	Jilin/Liaoning/Heilongjiang	62	0.028	0.435	0.124	0.097
Cold browncalcic soils	Tibet	41	0.041	0.277	0.123	0.046
Alluvial soils	Jilin	11	0.04	0.152	0.117	0.037
boggy soils	Hebei/Jilin/Tianjin	3	0.085	0.143	0.116	0.029
Purplish soils	Sichuan	56	0.037	0.269	0.11	0.055
Limestone soils	Yunnan/ Henan	33	0.028	0.25	0.108	0.076
Yellowbrown soils	Henan	7	0.083	0.123	0.1	0.016
Anthropogenic-alluvial soil	Henan	1	0.097	0.097	0.097	0
Cinnamon soils	Hebei/ Henan / Liaoning /Shandong/Shanxi	116	0.031	0.349	0.092	0.038
Fluvo-aquic soils	Henan/Hebei/Liaoning/Shandong/Tianjin/Shanxi	427	0.028	0.38	0.088	0.034
Latosols	Hainan	25	0.048	0.142	0.086	0.031
Paddy soils	Sichuan/Xinjiang/Zhejiang	271	0.016	0.269	0.074	0.038
Total	—	1661	0.016	0.461	0.113	0.062

仪预热 1 h 后在暗室条件下进行测定。每个土样采集 10 条光谱曲线,算术平均后作为该土样的实际反射率。

测得的光谱曲线首尾两端噪声明显,剔除后保留 400~2450 nm 范围波段。然后采用一阶微分(1std)配合 Savitzky-Golay 平滑(SG)进行预处理。为压缩土壤高光谱数据,采用主成分分析法(PCA)对预处理后的光谱进行降维。本文所有数据处理过程均由 Unscrambler X10.1,R 和 FuzzyME 实现。

2.3 建模方法及评价指标

2.3.1 偏最小二乘回归

偏最小二乘回归(PLSR)是一种基于因子分析的非参数回归方法,在光谱分析中最为经典和常用。该方法集成了主成分分析、典型相关分析和线性回归分析的优点,实现了数据结构的简化,解决了自变量之间多重共线性问题,适合用于光谱分析这种自变量较多的情况^[23]。

2.3.2 局部加权回归

局部加权回归(LWR)方法的思想是在光谱库中搜索和预测样本光谱空间距离最接近的若干样本,然后使用这些样本来拟合一个局部的线性回归方程^[24]。

LWR 算法首先确定用于每个局部建模集的主成分数 A 和样本数 C 。再对所有建模集进行 PCA

分析。 T 定义为得分矩阵, P 定义为载荷矩阵。对每一个预测样本计算得分 $t^T = x^T P$,其中 x 是通过与建模矩阵 X 相同的常数中心化得到的预测光谱。找到最接近预测样本的 C 个建模样本。接近程度通过 A 个尺度中主成分得分的空间距离测量来定义。计算 A 个尺度得分向量和 $T = XP$ 所在的所有行之间的距离,然后找到最接近预测样本的 C 个样本。

对于模型,可用最小二乘准则用来评估未知参数,再预估预测样本的全氮含量。LWR 中的一些调整参数是需要确定的,主要包括以下几个参数:样本间相似度计算的选择;加权最小二乘回归中加权函数的选择;在每个局部回归中主成分数 A 和样本数 C 的选择。

2.3.3 模糊 K 均值聚类

模糊 K 均值聚类(FKMC),基本思想是将一个数据集分成 k 个类别,寻找目标函数的迭代最小化,而其优势是能给出最佳分类数的分析指标,正适合应用于大样本光谱库分类。FKMC 方法的目标函数^[25]为

$$J_{\varphi}(\mathbf{M}, \mathbf{C}) = \sum_{i=1}^n \sum_{j=1}^k \mu_{ij}^{\varphi} d_{ij}^2, \quad (1)$$

式中 n 是属性的个数,此处为研究所选取的主成分个数, k 是类别个数, μ_{ij} ($1 \leq i \leq n, 1 \leq j \leq k$) 表示矩阵 \mathbf{X} 中第 i 个土壤样本 x_i 属于聚类中心矩阵 \mathbf{C} 中

第 j 个聚类中心 c_j 的隶属度, \mathbf{M} 为隶属度矩阵, d_{ij}^2 等于 x_i 与 c_j 在特征向量上距离的平方, φ 为模糊加权指数 ($\varphi \geq 1$), 它控制了不同类别间共用数据的数目。引入模糊性能指数 (FPI)、改进分类熵 (MPE) 和聚类独立性指标 S 来确定合适的聚类数, 作为最佳光谱分类数。这三类指标越小, 分类效果越好^[26]。

2.3.4 评价指标

预测模型利用决定系数 V_{R^2} 、均方根误差 V_{RMSE} 、测定值标准偏差与标准预测误差的比值 V_{RPD} 以及四分位数间距 ($Q_3 - Q_1$, 其中 Q_1 和 Q_3 分别为第一和第三四分位数) 与 V_{RMSE} 的比值 V_{RPIQ} 为精度评价指标。 V_{R^2} 、 V_{RPD} 及 V_{RPIQ} 越大, V_{RMSE} 越小, 说明预测效果越好。另外, 当 $1.5 < V_{RPD} < 2$ 时, 表明模型只能进行粗略估测, 当 $2.0 < V_{RPD} < 2.5$ 时, 表明模型具有较好的定量预测能力, 当 $2.5 < V_{RPD} < 3.0$ 时, 模型具有很好的预测能力, 当 $V_{RPD} > 3.0$ 时模型具有极好的预测能力。Bellon-Maurel 等^[27] 认为基于四分位差 ($Q_3 - Q_1$), 使预测误差标准化的评价指标 V_{RPIQ} 对于预测结果偏态分布的描述比指标 V_{RPD} 更稳健合理, V_{RPIQ} 值越大, 表明预测结果分布的偏态程度越小。

3 结果与讨论

3.1 土壤光谱库 PLSR 全局建模预测 TN 含量

将预处理后 1661 个库中样本的光谱数据 (400~2450 nm) 与对应的 TN 值进行 PLSR 全局建模, 模型采用留一法交叉验证, 经比较设定因子数为 15 最合适, 可避免过度拟合。用建立好的全局模型预测 104 个浙江水稻土样本, 并与对应的 TN 真实值进行最终检验, 反演结果如图 1 所示, P_1 表示 PLSR 模型, $V_{R^2}^{P_1}$ 表示该样本 PLSR 的预测模型决定系数, $V_{RMSE}^{P_1}$ 为该样本 PLSR 的均方根误差, $V_{RPD}^{P_1}$ 为该样本 PLSR 标准偏差与标准预测误差的比值, $V_{RPIQ}^{P_1}$ 为该样本 PLSR 四分位数间距与均方根误差的比值。

从图 1 可以看出, TN 实测值大于 0.2% 的水稻土样本, 预测结果普遍存在低估现象, 预测样本点逐渐偏离 1:1 线, 呈现与坐标横轴近似平行走向的趋势。

经统计, 本研究建立的库中所有样本 TN 值大于 0.2% 的样本数目占全库样本的 10.05%, PLSR 全局建模时贡献的信息量占库信息总量比例偏少, 在预测高 TN 含量水稻土样本时易出现低估现象, 不能充分挖掘其预测能力, 这是造成整体预测精度

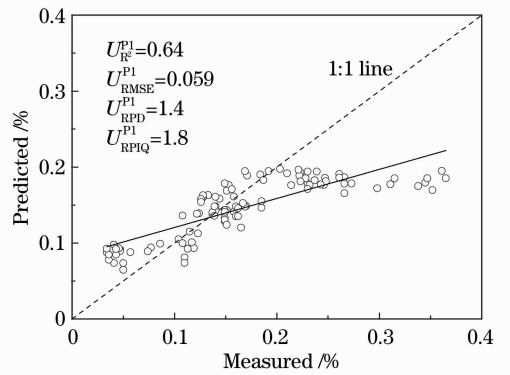


图 1 PLSR 全局建模预测结果

Fig. 1 Prediction result of PLSR model

偏低的原因。

3.2 土壤光谱库 LWR 局部建模预测 TN 含量

采用 LWR 方法, 光谱库局部建模预测未知样本 TN 时, 对每一个未知样本, 需要从库中挑选出若干光谱空间距离较近的样本点。然后将挑选出的样本和未知样本光谱数据一起进行主成分分析, 选取若干个主分变量回归分析, 来预测未知样本的 TN 含量, 目的是降维, 提取有效光谱信息。所以, 从库中挑选的样本点和主分变量的个数是 LWR 方法的关键参数。

参考 Ramirez-Lopez 等^[28] 2013 年的研究结果, 经过不同参数组合, 对比后发现, 将参数设定为 50 个相似样本点和 5 个主成分个数得到的预测精度最高。通过计算某一未知样点在空间马氏距离中与其最近的库中样本点, 筛选出 50 个相似样本进行建模, 进行回归预测。同样的方法应用于所有未知样本。图 2 为 104 个水稻土样本实测预测结果, P_2 表示 LWR 模型, $V_{R^2}^{P_2}$ 表示该样本 LWR 的预测模型决定系数, $V_{RMSE}^{P_2}$ 为该样本 LWR 的均方根误差, $V_{RPD}^{P_2}$ 为该样本 LWR 的测定标准差和标准预测误差的比值, $V_{RPIQ}^{P_2}$ 为该样本 LWR 的四分位数间距与均方根

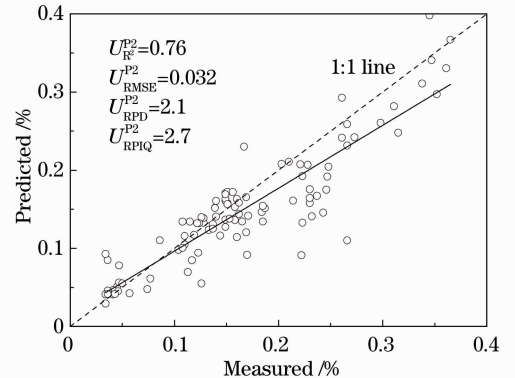


图 2 LWR 局部建模预测结果

Fig. 2 Prediction result of LWR model

误差的比值。LWR 方法能够根据光谱样本空间马氏距离,从库中选出相似样本,利用库中有效信息建模预测,整体反演精度显著提高。模型具有很高的稳定性和普适性,为进一步探讨光谱库建模方法提供了依据。

图 3 为某一未知水稻土样本从光谱库中搜索临近样点的样本在光谱主成分空间的分布情况,前两个主分 PC1 和 PC2 占解释变量信息的 71%。由图中可以清楚看出,依据未知样本与库中样本空间马氏距离的远近,挑选出的建模样本集中紧密分布在未知样本周围。这样,在建立回归模型时,剔除了库中的干扰信息,选取的相似样本有效信息能更好预测未知样本 TN 含量。

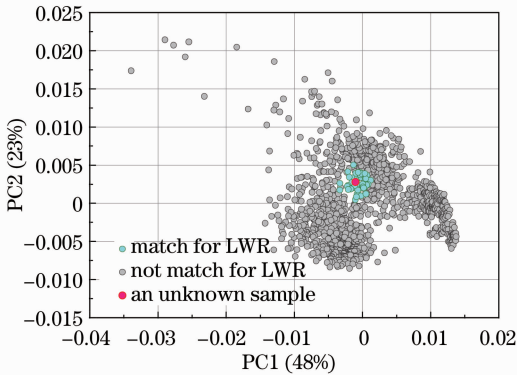


图 3 某待测样本 LWR 匹配样点在主成分 PC1 和 PC2 空间的分布

Fig. 3 Distributions of one unknown sample matched by LWR in PC1 and PC2 spaces

3.3 土壤光谱库 FKMC-PLSR 局部建模预测 TN 含量

可先对所有预处理后的光谱数据降维,取前三个因子(PCA1, 48%; PCA2, 23%; PCA3, 7%; 累积贡献率达 78%)作为聚类输入数据。FKMC-PLSR 方法需根据参数模糊性能指数、改进分类熵和独立模糊校验指数的值来确定最佳分类数目,三者均是越小越好。为了找出最佳分类数目,分别产生 2、3、4、5、6、7、8、9 个类别。图 4 是将所有土壤按照反射光谱特性划分为 2、3、4、5、6、7、8、9 时得到的 FPI、MPE 和 S 的值。可以看出,当分类数目为 5 时, FPI、MPE 和 S 的值均为最小,最佳的分类数目为 5 类。按照最终 5 类聚类结果,将库中 1661 样本分别 PLSR 建模,并预测归属到对应类别中的未知样本 TN 值。104 个样本的预测结果如图 5 所示, P3 表示 FKMC-PLSR 模型, $V_{R^2}^P$ 为样本 FKMC-PLSR 的预测模型决定系数, V_{RMSE}^P 为样本 FKMC-

PLSR 的均方根误差, V_{RPD}^P 为 FKMC-PLSR 的测定标准差和标准预测误差的比值, V_{RPIQ}^P 为样本 FKMC-PLSR 的四分位数间距和均方根误差的比值。

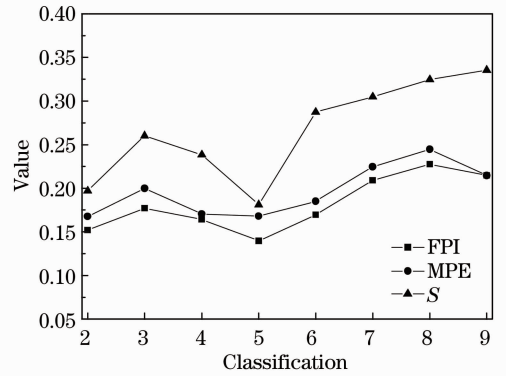


图 4 FKMC-PLSR 分成不同类别时的 FPI、MPE 和 S 值
Fig. 4 FPI, MPE and S value versus number of clusters

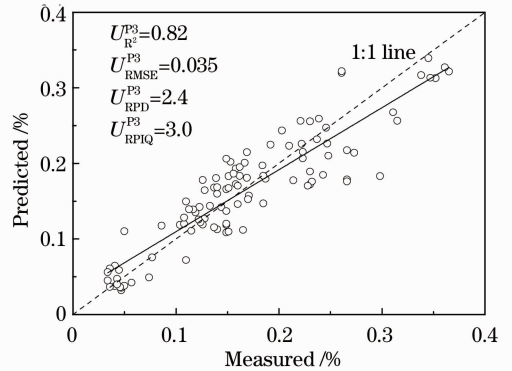


图 5 FKMC-PLSR 局部建模预测结果

Fig. 5 Prediction result of FKMC-PLSR model

3.4 三种土壤 TN 含量反演模型的对比

分别使用 PLSR 全局、LWR 和 FKMC-PLSR 局部三种方法对库建模,来预测浙江 104 个水稻土样本 TN 含量。三种预测模型精度指标对比如表 2 所示。从表 2 可以看出,LWR 和 FKMC-PLSR 局部模型的预测精度均优于 PLSR 全局模型,并且提高显著。

表 2 不同模型预测结果精度指标结果对比
Table 2 Comparison of prediction accuracy by different models

Model	Prediction accuracy			
	V_{R^2}	V_{RMSE}	V_{RPD}	V_{RPIQ}
PLSR	0.64	0.059	1.4	1.8
LWR	0.76	0.032	2.1	2.7
FKMC-PLSR	0.82	0.035	2.4	3.0

PLSR 全局建立的模型,对于预测每一个样本,都包含有库中的有效信息和干扰信息,没有兼顾不同样本光谱曲线形状的差异。在利用光谱库预测局

部区域 TN 含量时,该模型反演高 TN 含量土样能力不足,存在低估现象,导致整体反演精度较低。

LWR 和 FKMC-PLSR 方法均是从库中挑选相似光谱曲线的样本,构成数据子集进行局部建模。因为光谱库中信息量丰富,很容易从中挑选出光谱空间距离相近的土样,充分挖掘库中有效信息,解决了全局 PLSR 模型高 TN 含量样本反演能力弱的缺点,预测模型具有很好的稳定性和普适性,预测能力显著提高。但两种方法挑选建模子集的方式不同:LWR 方法对每一待预测样本都要挑选一定量的相似样本,建模子集之间可能有重复样本,工作量大;FKMC-PLSR 直接基于最佳聚类类别划分库为若干子集分别建模,子集之间没有交集,待预测样本依据所属类别利用相应模型进行预测。同时,LWR 只能依据光谱空间位置来构建子集,FKMC-PLSR 还可将聚类的依据扩展到土壤属性。

4 结 论

研究建立的大样本土壤光谱库地域分布广阔,成土母质差异性较大,包含信息量丰富。针对不同地区,建立普适预测模型充分挖掘有效信息是光谱库共享的必要前提。比较三种对库建立的预测模型,得出以下结论:

1) 当前,基于全国不同土壤类型的大样本光谱库建立一个通用预测模型尚不可行。PLSR 全局模型将库中将所有信息用来预测,难以有效针对每一待测样本从库中挖掘有效信息。

2) 依据待预测样本光谱信息,从库中挑选相似样本进行局部建模的 LWR 和 FKMC-PLSR 方法能够充分挖掘库中有效信息。较 PLSR 全局模型预测精度显著提高, V_{R^2} 由 0.64 分别迅速提升到 0.76 和 0.82,提升幅度分别达 18.75% 和 28.13%。两种方法 V_{RPD} 分别为 2.1 和 2.4,均大于 2,模型预测能力显著增强,可以进行 TN 含量精确定量的研究。

3) LWR 和 FKMC-PLSR 模型具有很好的稳定性和普适性,光谱库中并未包含浙江水稻土,但在预测时空差异较大的区域样本时得到很好的反演结果。

提出从全国光谱库中局部优选建模的 LWR 方法和将库按类划分的 FKMC-PLSR 方法,为如何充分挖掘光谱库有效信息、建立普适性强的 TN 含量预测模型提供了可靠依据。

参 考 文 献

1 A V Bilgili, H M V Es, F Akbas, *et al.*. Visible-near infrared

reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey [J]. *Journal of Arid Environments*, 2010, 74(2): 229–238.

2 Lu Wanzhen. *The Modern Analysis Technique of Near-Infrared Spectrum* (2nd edition) [M]. Beijing: China Petrochemical Press, 2006. 10–13.

陆婉珍. 现代近红外光谱分析技术(第2版)[M]. 北京: 中国石化出版社, 2006. 10–13.

3 Chen Cong, Lu Qipeng, Peng Zhongqi. Preprocessing methods of near-infrared spectrum based on NLMS adaptive filtering [J]. *Acta Optica Sinica*, 2012, 32(5): 0530001.

陈 丛, 卢启鹏, 彭忠琦. 基于 NLMS 自适应滤波的近红外光谱去噪处理方法研究[J]. *光学学报*, 2012, 32(5): 0530001.

4 Gao Hongzhi, Lu Qipeng, Ding Haiquan, *et al.*. Robust calibration methods of near-infrared spectrum based on random sample consensus algorithm [J]. *Acta Optica Sinica*, 2013, 33(s2): s230001.

高洪智, 卢启鹏, 丁海泉, 等. 基于随机抽样一致性算法的近红外光谱稳健模型研究[J]. *光学学报*, 2013, 33(s2): s230001.

5 B Stenberg, R A V Rossel, A M Mouazen, *et al.*. Visible and near infrared spectroscopy in soil science [J]. *Advances in Agronomy*, 2010, 107: 163–215.

6 B Y Kuang, H S Mahmood, M Z Quraishi, *et al.*. Sensing soil properties in the laboratory, *in situ*, and on-line: a review [J]. *Advances in Agronomy*, 2012, 114: 155–223.

7 Y X Song, F L Li, Z F Yang, *et al.*. Diffuse reflectance spectroscopy for monitoring potentially toxic elements in the agricultural soils of Changjiang River Delta, China [J]. *Applied Clay Science*, 2012, 64: 75–83.

8 M Nocita, A Stevens, G Toth, *et al.*. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach [J]. *Soil Biology and Biochemistry*, 2014, 68: 337–347.

9 D J Brown, K D Shepherd, M G Walse, *et al.*. Global soil characterization with VNIR diffuse reflectance spectroscopy [J]. *Geoderma*, 2006, 132(3-4): 273–290.

10 R A V Rossel, R Webster. Predicting soil properties from the Australian soil visible-near infrared spectroscopic database [J]. *European Journal of Soil Science*, 2012, 63(6): 848–860.

11 Z Shi, Q L Wang, J Peng, *et al.*. Development of national VNIR soil-spectral library for soil classification and the predictions of organic matter [J]. *Science China Earth Sciences*, 2014, 57(7): 1671–1680.

12 D J Brow. Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed [J]. *Geoderma*, 2007, 140(4): 444–453.

13 J Wetterlind, B Stenberg. Near-infrared spectroscopy for within-field soil characterization: small local calibrations compared with national libraries spiked with local samples [J]. *European Journal of Soil Science*, 2010, 61(6): 823–843.

14 F Goge, R Joffre, C Jolivet, *et al.*. Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database [J]. *Chemometr Intell Lab*, 2012, 110(1): 168–176.

15 A Stevens, M Nocita, G Toth, *et al.*. Prediction of soil organic carbon at the european scale by visible and near infrared reflectance spectroscopy [J]. *Plos One*, 2013, 8(6): 1–13.

16 Y Peng, M Knadel, R Gislum, *et al.*. Predicting soil organic carbon at field scale using a national soil spectral library [J]. *J Near Infrared Spectroscopy*, 2013, 21(3): 213–222.

17 Xu Yongming, Lin Qizhong, Huang Xiuhua, *et al.*. Experimental study on total nitrogen concentration in soil by VNIR reflectance spectrum [J]. *Geography and Geo-Information Science*, 2005, 21(1): 19–22.

徐永明, 蔺启忠, 黄秀华, 等. 利用可见光/近红外反射光谱估算

- 土壤总氮含量的实验研究[J]. 地理与地理信息科学, 2005, 21(1): 19-22.
- 18 Lu Yanli, Bai Youlu, Wang Lei, *et al.*. Determination for total nitrogen content in black soil using hyperspectral data [J]. Transactions of the CSAE, 2010, 26(1): 256-261.
- 卢艳丽, 自由路, 王磊, 等. 黑土土壤中 TN 含量的高光谱预测分析[J]. 农业工程学报, 2010, 26(1): 256-261.
- 19 Li Shuo, Wang Shanqin, Zhang Meiqin. Comparison among principal component regression, partial least squares regression and back propagation neural network for prediction of soil nitrogen with visible-near infrared spectroscopy [J]. Acta Optica Sinica, 2012, 32(8): 0830001.
- 李硕, 汪善勤, 张美琴. 基于可见近红外光谱比较主成分回归、偏最小二乘回归和反向传播神经网络对土壤氮的预测研究[J]. 光学学报, 2012, 32(8): 0830001.
- 20 Zhang Juanjuan, Tian Yongchao, Yao Xia, *et al.*. Estimating soil total nitrogen content based on hyperspectral analysis technology [J]. J Natural Resources, 2011, 26(5): 881-890.
- 张娟娟, 田永超, 姚霞, 等. 基于高光谱的土壤 TN 含量估测[J]. 自然资源学报, 2011, 26(5): 881-890.
- 21 Zhang Juanjuan, Tian Yongchao, Yao Xia, *et al.*. Estimating model of soil total nitrogen content based on near-infrared spectroscopy analysis [J]. Transactions of the CSAE, 2012, 28(12): 183-188.
- 张娟娟, 田永超, 姚霞, 等. 基于近红外光谱的土壤 TN 含量估算模型[J]. 农业工程学报, 2012, 28(12): 183-188.
- 22 Ji Wenjun, Shi Zhou, Zhou Qing, *et al.*. VIS-NIR reflectance spectroscopy of the organic matter in several types of soils [J]. Journal of Infrared and Millimeter Waves, 2012, 31(3): 277-282.
- 纪文君, 史舟, 周清, 等. 几种不同类型土壤 VIS-NIR 光谱特性及有机质响应波段[J]. 红外与毫米波学报, 2012, 31(3): 277-282.
- 23 Ji Wenjun, Li Xi, Li Chengxue, *et al.*. Using different data mining algorithms to predict soil organic matter based on visible-near infrared spectroscopy [J]. Spectroscopy and Spectral Analysis, 2012, 32(9): 2393-2398.
- 纪文君, 李曦, 李成学, 等. 基于全谱数据挖掘技术的土壤有机质高光谱预测建模研究[J]. 光谱学与光谱分析, 2012, 32(9): 2393-2398.
- 24 T Naes, T Isaksson, B Kowalski. Locally weighted regression and scatter correction for near infrared reflectance data [J]. Analytical Chemistry, 1990, 62(7): 664-673.
- 25 Y Guo, Z Shi, H Y Li, *et al.*. Application of digital soil mapping methods to identify salinity management classes in coastal lands of central China [J]. Soil Use and Management, 2013, 29(3): 445-456.
- 26 Y Li, Z Shi, F Li, *et al.*. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land [J]. Computers and Electronics in Agriculture, 2007, 56(2): 174-186.
- 27 V Bellon-Maurel, E Fernandez-Ahumada, B Palagos, *et al.*. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy [J]. TrAC Trends in Analytical Chemistry, 2010, 29(9): 1073-1081.
- 28 L Ramirez-Lopez, T Behrens, K Schmidt, *et al.*. Distance and similarity-search metrics for use with soil vis-NIR spectra [J]. Geoderma, 2013, 199: 43-53.

栏目编辑：史敏